图书情报工作
第65卷 第5期 2021年3月

"十三五"图情档学科进展:外文期刊论文解析*

赵蓉英[1,2,3] 朱伟杰[1,3] 常茹茹[1,3] 段香廷[1,3]

[1] 武汉大学中国科学评价研究中心 武汉 430072 [2] 武汉大学信息资源研究中心 武汉 430072

[3] 武汉大学信息管理学院 武汉 430072

摘 要:[目的/意义]对图书情报领域的国际发文情况、研究热点以及学科前沿进行研究,有助于推动图情学科的发展,提升我国学术话语权。[方法/过程]以SCI和SSCI收录的图书情报领域高影响力外文期刊论文为数据源,利用HistCite、CiteSpace、SPSS等工具从研究现状、高影响力学者、高被引论文、研究热点、前沿趋势5个方面分析近5年来(2016-2020年)图情领域的发展状况。通过对高频关键词的年度变化分析,总结四种呈现不同态势的关键词,并对低频关键词进行透视,得出人文、技术和管理三个维度的发展主题。[结果/结论]随着信息技术的不断发展和信息环境的不断变化,图情学科研究主题由传统信息服务向知识服务,由知识组织与管理相关研究向知识组织体系构建等研究主题进一步延伸和细化,且注重理论和实践相结合已日渐成为图情档学科研究的重要趋势。

关键词:"十三五" 图情档 学科进展 外文期刊论文 词频分析

分类号:G250

DOI:10.13266/j.issn.0252-3116.2021.05.003

1 引言

把握领域知识结构与演变,正确评价基础科学研究已引起学界广泛关注,成为信息时代相关学者关注的重要研究课题。世界著名科学计量学家、普赖斯奖获得者H. D. White[1]教授在1998年就利用作者共被引分析法对图书情报学领域的12种期刊在1972-1995年间所刊载的学术论文进行了计量分析,对学科的现状、发展轨迹和趋势进行了系统的梳理和研究。回顾学科研究成果,以深入揭示学科发展轨迹;梳理学科研究现状,以更好地把握本领域研究动态[2];追踪学科研究热点前沿,以提升我国的学术话语权和图情学科研究水平。现有国内外图情学科进展研究论文均从定量分析与定性综述两种研究范式出发,从回顾、展望等角度揭示研究发现与结论,例如邱均平[2]等利用文献计量的方法对13种高影响力的外文源刊进行了分析,发现国际范围内图书情报学研究普遍从计量、管理、技术三个维度开展,领域研究将进入新一轮的繁荣期。鉴于此,本文以SCI和SSCI收录的高影响力外文源刊在2016-2020年所刊载的论文为研究对象,引入CiteSpace、HistCite和SPSS等工具,从文献统计、高影响力作者、高频关键词、低频关键词等方面对图书情报学领域的学术研究成果和研究热点进行了探讨,以期从科学量化角度揭示图情档学科的发展和特点。

2 数据来源与方法

本文以美国科学信息研究所(Institute for Scientific Information)出版的Web of Science数据库为数据源,该数据库具有收录文献时间跨度大、数量多、质量高等优势,能够最大限度地保证数据的完整性,同时检索结果的数据输出格式能被公共分析软件所接受。在数据库选择上,数据库的不同往往对结果会产生一定差异,宏观层面上而言,学界常用于文献计量学的指标都来自多学科的大型数据库,如SCI、CSCD等,此类数据库均试图覆盖全国或者是世界各国的文献。尽管不同数据库都会有所差异,每种文献计量学指标用于评价都有不完善之处,但在宏观的评价层次上数据库差异可忽略不计。微观层面上而言,数据库检索结果数量的差异也是直接影响综合分析的重要因素,以LISA数据库为例,以同样的检索方式进行文献统计,检索得到文献

* 本文系国家社会科学基金重大项目"构建中国话语权的评价科学理论、方法与应用体系研究"(项目编号:18ZDA325)研究成果之一。

作者简介:赵蓉英(ORCID:0000-0002-4742-9037),教授,博士;朱伟杰(ORCID:0000-0003-2970-5728),硕士研究生,通讯作者,E-mail:zhuwjzzz@163.com;常茹茹(ORCID:0000-0002-0607-7194),博士研究生;段香廷(ORCID:0000-0003-4368-9613),硕士研究生。

收稿日期:2020-07-19 修回日期:2020-10-15 本文起止页码:26-39 本文责任编辑:杜杏叶

5 239篇，由于 LISA 收录范围广，对具体文献的质量没有细致化的划分，根据帕累托关键少数法则，尽管选择 SCI-EXPANDED、SSCI 数据库所收录的文献相比较少，但能反映领域内的重要进展与趋势，考虑到 WoS 核心合集下 SCI-EXPANDED、SSCI 数据库的权威性和高质量，本研究选择这两大引文索引库作为数据检索来源。

具体方法为：以"WC =（Information Science & Library Science） AND CU =（China）"为检索式（该类别中含部分管理科学与工程专业相关期刊 Mis Quarterly、Information Systems Research、International Journal of Information Management 等），限定文献收录时间为 2016－2020 年，引文数据库为 SSCI、SCI－EXPANDED，文献类型为期刊论文（Article），精炼检索后的有效文献数据为 2 270 篇（截至 2020 年 4 月 25 日）。对 2 270 篇的原始文献数据进行导出，输出记录内容设定为全记录（包含引用的参考文献），并将文件格式以纯文本保存。

本文主要采用美国 Drexel 大学陈超美教授开发的 CiteSpace 与 IBM 旗下的统计软件 SPSS 进行词频统计与可视化分析，借助引文分析软件 HistCite 分析论文在 Web of Science 数据库中的被引频次和论文在所收集的样本数据集中的被引频次。运用 CiteSpace5.5.R2、SPSS v26.0 和 HistCite Pro 2.1 版本，在对不同参数进行调整的基础上，采用高频关键词年度变化分析、低频关键词分类等方法对研究热点和前沿进行分析，通过高频关键词年度变化来确定研究热点，通过低频关键词分类分析图情领域的前沿趋势，并综合这几个方面分析领域的发展状况。

3 结果分析

3.1 文献统计

文献信息量统计是科学基础理论研究的一个重要环节，一定程度上反映出该学科领域的研究热度，图情领域外文期刊 2016－2020 年发表文献共 2 270 篇（数据统计截至 2020 年 4 月 25 日），"十三五"期间年均发文超过 400 篇，且从 2018 年开始有了明显的增长趋势（2020 年数据不完整，暂不做详细分析）。

从引文情况来看，2 270 篇论文被引频次总计 12 522次，每篇平均引用次数为 5.79。"十三五"期间，图情学科研究在广度和深度上都取得了开拓性进展，研究成果不断丰富，涵盖多种研究类型的期刊论文，为进一步统计论文在当前数据库中被引用的次数与全球关注度，本文引入文献分析工具 Histcite 的两个引文指标 LCS（Local Citation Score）和 GCS（Global Citation Score）进行分析。通过 LCS 间接反映领域内的研究状况，通过 GCS 间接反映学科对科学研究的影响，相关引文数情况如表 1 所示：

表 1 "十三五"期间外文期刊论文基本数据

年份	P（发文篇数）	Q（被引次数）	LCS（Q）	GCS（Q）	LCS（P）	GCS（P）
2016	403	5 142	589	4 884	1.46	12.12
2017	451	4 154	565	3 987	1.25	8.84
2018	498	2 309	391	2 241	0.79	4.50
2019	679	825	111	798	0.16	1.18
2020	239	92	8	92	0.03	0.38

由于论文引用会出现时滞现象，且引文峰值一般在论文发表 2 年以后，因此"十三五"期间的 LCS 值和 GCS 值是下降的，从整体的趋势来看：图情领域发文量不断增长，近 15 年来的研究成果增速十分明显（见图 1）。从 2006 年开始至今，我国图书情报学科外文期刊论文发文数逐年递增，"十三五"期间所发表论文数大于前十年论文发表数量的总和，且在 2019 年达到增幅的最高值，2020 年由于只统计至 4 月，文献数较少。LCS 在 2013 年达到顶峰，而 GCS 最高点则是在 2012 年，此前 LCS 和 GCS 值也处于一个较低的水平上，这也反映出了当时由于自身研究水平不高，不能对外部领域形成较明显的影响，而后随着新兴信息科学技术的普及，与其他学科的联系、融合也更为密切。综合图 1 与图 2 可知，图书情报领域目前正处于开放增长期的发展阶段，图情学科国际发文快速增长，学术成果更注重国际影响力，这同时也说明了图情学科的研究投入有了明显的增加。

3.2 高发文、高被引作者分析

高影响力学者不仅包括学科、流派的缔造者和领域经典理论的奠基者，还包括时下活跃在科研一线的研究人员[3]，对论文高产出与高被引作者进行统计分析，可以了解研究领域的核心人物与学科科研实践的趋向[4]。通过对作者发文量和被引频次的统计分析，得出"十三五"期间图情领域外文期刊中高影响力作者的分布情况，统计结果表明，十三五期间图情领域外文期刊发文数量最多的有 12 篇，被引频次最高的有 210 次。

根据普赖斯定律，半数的论文为一群高生产能力作者所撰，这一作者集合的数量上约等于全部作者总数的平方根[5]，假设最高产的那位学者所发表的论文数为 n_{max}，将所有发表论文的总数记为 $P(1,n_{max})$，则普赖斯定律可用公式(1)表示[6]。

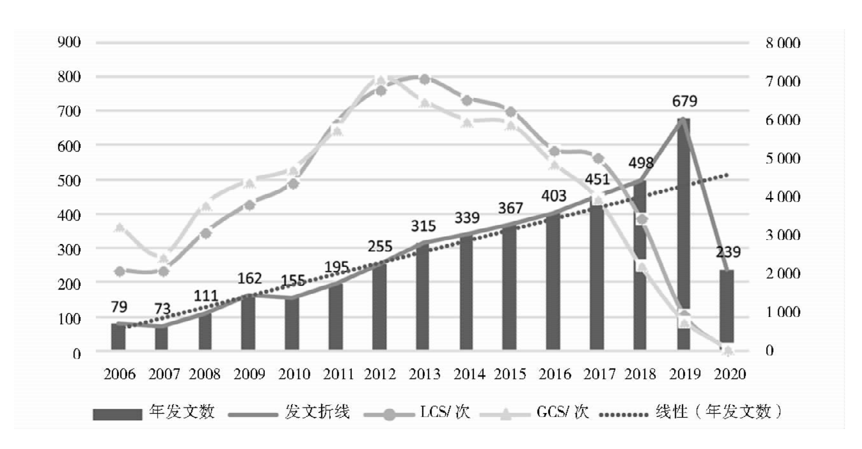


图 1　2006 年 – 2020 年图情档学科外文期刊论文历年发文与引用情况一览

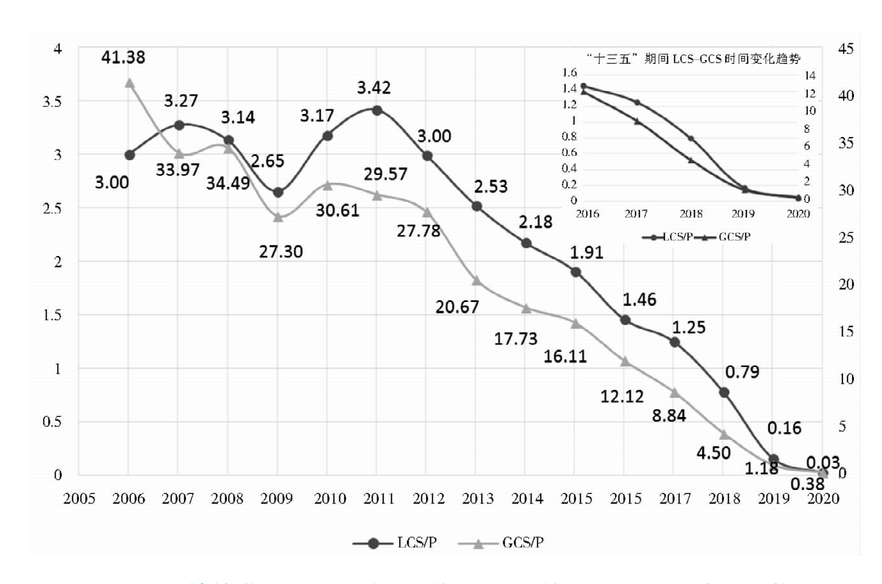


图 2　图情档学科外文期刊论文篇均 LCS、篇均 GCS 时间变化趋势

$$\frac{1}{2}P(1,m_{max}) = P(m,n_{max}) = P(1,m) \qquad 公式(1)$$

其中 m 为普赖斯所假定的一个数，即个人论文数大于 m 的科学家们所发表的论文总数等于全部论文总数的一半，在 m 值的确定上，普赖斯根据洛特卡定律，借用数学结论得到 $m \approx 0.749(n_{max}^{1/2})$[7]，即论文产量大于 3 的作者可以界定为高产作者，则图书情报学科外文期刊中高产作者共有 124 人，发文数量排名前 5 的作者分别为吴丹、P. Lo、D. Nicholas、赵蓉英和胡晓，其中 P. Lo 和 D. Nicholas 与中国学者合作度较高，研究领域相对成熟。高产的核心作者研究工作较为连续，研究成果对于图情学科领域后续研究具有较大的参考价值。此外，高产作者之间往往具有关联性，研究领域也多有交叉，学者间学术交流与科研合作增强。引频次排名前 5 的作者分别为 V. Venkatesh、W. L. Shiau、I. Hashem、T. Abaker、Y. Yao 和 S. Nambisan，他们是图情学科的高影响力作者，为图书馆学情报学的发展作出了重要贡献，在图情界发挥着学术引领的作用。此外，张琳、刘晓萍、P. B. Lowry 等学者在图情领域也取得了丰硕的研究成果，具有一定的影响力。就高被引论文而言，作为体现每个学者对国家、机构的学术贡献度，以及国家/地区、机构学术声誉重要标志领域高被引论文中被引前三的均是国外学者，国内学者今后可以加强国际交流，在信息资源管理、信息计量等领域选择国外顶尖研究机构（如莱顿大学等）进行合作，产出更多的高被引论文，详见图 3、图 4。

3.3　外文期刊高被引论文分析

由于科学研究论文的引文峰值一般为 2 年以上，本次所统计的文献最近发表时间为 2020 年，因此 2019 年与 2020 年的被引数据没有实际意义，在此不作分析。表 2 是 2016 年、2017 年、2018 年图情学科外文期刊中被引排序前 5 的论文。综合分析表 2 的高被引论文，可以了解该时间段下图书情报学领域的关注热点，其主题集中在与图书情报学有关的学术研究、技术研究和大数据时代下相关社会关系讨论等方面。

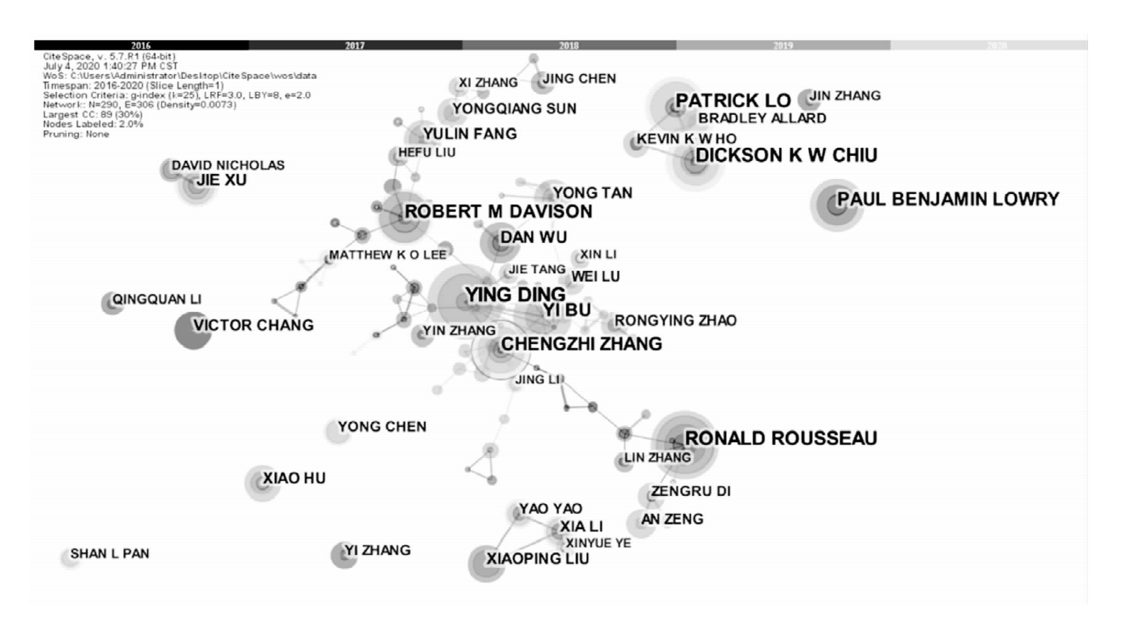


图 3 “十三五”期间图情档学科外文期刊作者合作网络图谱

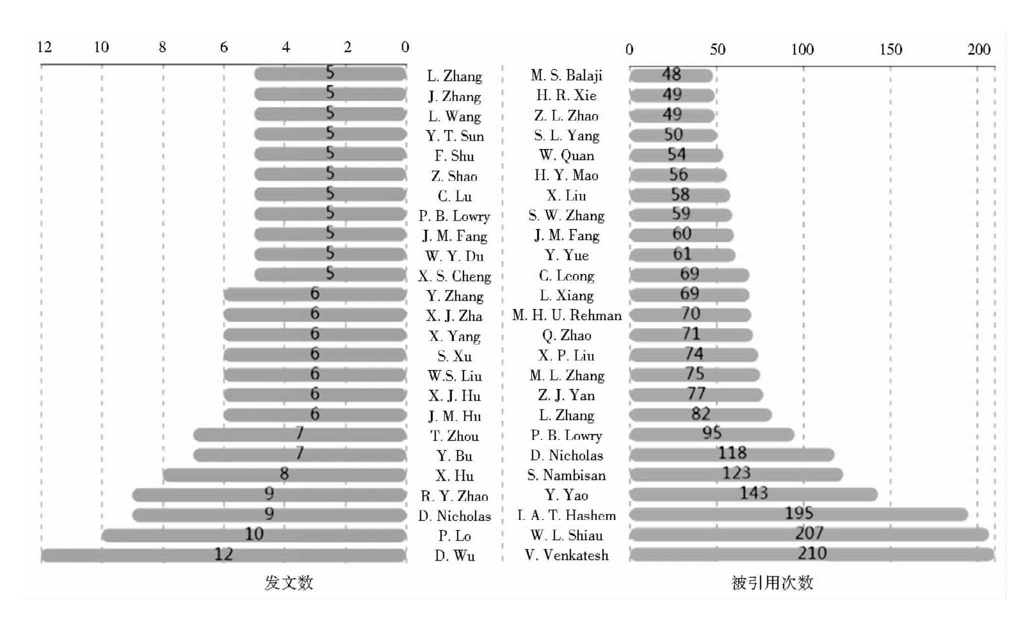


图 4 “十三五”期间图情档学科外文期刊高发文、高被引作者一览

在图书情报学相关理论研究方面，*Digital innovation management: Reinventing innovation management research in a digital world*[8] 旨在推动新数字创新管理的新理论，对创新的定义边界、创新机构以及创新过程和结果之间的关系等基本假设提出了质疑；*Emerging trends and new developments in information science: A document co-citation analysis* (2009－2016)[9] 则描述了知识的结构、研究课题的演变和课题的出现一直是信息科学的重要组成部分。

相关的技术研究方面，*The role of big data in smart city*、*Understanding behavioral intention to use a cloud computing classroom: A multiple model comparison approach*、*Sensing spatial distribution of urban land use by integrating points-of-interest and Google Word2Vec model*[10－12] 等研究了在信息化社会不断发展的基础下，运用云计算等信息技术将收集的信息资源进行高效利用，促进智慧城市的发展或者城市土地的运用。

此外，大数据时代下相关社会关系方面的外文期刊论文较多，由此可推测研究者们注重图书情报对现实社会关系的影响。*Big data reduction framework for value creation in sustainable enterprises*[13] 构建了企业应用程序中进行端到端的数据缩减业务模型，并提出了一种新颖的概念－减少大数据在客户端中早期数据执行还原操作来实现多个目标，创造企业价值；*Digital*

transformation by SME entrepreneurs: A capability perspective[14]对阿里巴巴数字平台上的中小企业跨境电商的数字化转型进行了定性研究,调查能力不足和资源有限的中小企业企业家如何推动企业的数字化转型,提出数字创业和数字转型的新见解。这些研究成果都反映了图情学者研究视域趋于多样化,数据时代和智能时代的到来也引发社会学反思,隐私、风险等问题引起热烈讨论。

表 2　2016－2018 年图情档学科外文期刊高被引论文(部分)

论文题目	第一作者	期刊	卷期	被引用次数	年份
The role of big data in smart city	I. A. T. Hashem	*International Journal of Information Management*	36	193	2016
Unified theory of acceptance and use of technology: A synthesis and the road ahead	V. Venkatesh	*Journal of the Association for Information Systems*	17	143	2016
Knowledge sharing in online health communities: A social exchange theory perspective	Z. J. Yan	*Information & Management*	53	77	2016
Understanding behavioral intention to use a cloud computing classroom: A multiple model comparison approach	W. L. Shiau	*Information & Management*	53	74	2016
Big data reduction framework for value creation in sustainable enterprises	M. H. U. Rehman	*International Journal of Information Management*	36	71	2016
Digital innovation management: Reinventing innovation management research in a digital world	S. Nambisan	*Mis Quarterly*	41	123	2017
Sensing spatial distribution of urban land use by integrating points-of-interest and Google Word 2Vec model	Y. Yao	*International Journal of Geographical Information Science*	31	83	2017
Influence of customer engagement with company social networks on stickiness: Mediating effect of customer value creation	M. L. Zhang	*International Journal of Information Management*	37	75	2017
Co-citation and cluster analyses of extant literature on social networks	W. L. Shiau	*International Journal of Information Management*	37	73	2017
Measurements of POI-based mixed use and their relationships with neighbourhood vibrancy	Y. Yue	*International Journal of Geographical Information Science*	31	61	2017
Examining the core knowledge on facebook	W. L. Shiau	*International Journal of Information Management*	43	60	2018
The dynamics of ICT, foreign direct investment, globalization and economic growth: Panel estimation robust to heterogeneity and cross-sectional dependence	Z. Latif	*Telematics and Informatics*	35	38	2018
Digital transformation by SME entrepreneurs: A capability perspective	L. Li	*Information Systems Journal*	28	34	2018
Health information privacy concerns, antecedents, and information disclosure intention in online health communities	X. Zhang	*Information & Management*	55	33	2018
Creating strategic business value from big data analytics: A research framework	V. Grover	*Journal of Management Information Systems*	35	33	2018

3.4　热点分析

关键词分析的统计数据来源于 2016 到 2020 年发表的论文,不包括增刊和非学术性文章(2020 年期刊数据统计尚不完整)。经过抽取统计期刊论文关键词,共得到 7 280 条关键词数据,覆盖论文数共计 2 270 篇。

由于关键词属于自然语言,有些不能明确地反映研究主题,因此本文在统计时将检索意义不大的关键词列出,比如"China""Strategy""Motivation""Impact""Support"等,对此类词汇进行甄别、删除与规范。具体做法是:没有实际检索意义的关键词一律剔除;对单复数、缩写词、同义词等进行合并处理;出现同义词时,选择最常用的关键词为标准进行统计,其他相关词作补充;对有效关键词进行统一的术语规范处理与合并,比如"Wechat""Facebook""Twitter""Micro blog"合并为"Social media""Web""Network"合并为"Internet""University library""College library"合并为"Academic library"。在处理过程中,对同义词与近义词规范合并,最后对通用词进行统计、累加、排序。

由于关键词较多,本文主要采用普赖斯公式与孙清兰于 1992 年提出的高低频次界分方法综合选取词频阈值,根据普赖斯公式:

$$M = 0.749\sqrt{N_{max}} \qquad \text{公式(2)}$$

其中,N_{max}表示十三五期间学术论文被引频次的最高值,M 表示高频词阈值,由表 2 可知 $N_{max} = 195$,因此 M＝10(原 M 值为 10.459,在此取整),依据孙清兰所提出的界分公式:

$$n = \frac{1}{2}(-1 + \sqrt{1-4D}) \qquad 公式(3)$$

其中 D 表示不同词数[15]。在对关键词进行分类、筛选、合并规范后,最终得到的关键词数为 1031,由公式(3)可得高低频词界分值 n 为 31(原 n 值为 31.613,在此取整),两种高低频次划分方式在学界均被广泛采用,但本文中所计算差值较大,为避免单一取值倾向,综合两者所得出的值,本文取词频阈值为 20。

经过多轮筛选与数据处理,最终得到频次在 10 以上的关键词 223 个,为了更准确分析"十三五"期间高频关键词的变化趋势,本文还统计了高频词在"十一五"和"十二五"十年里所出现的频次,综合后可将"十三五"期间的高频关键词分为四类:第一类是在近 5 年尽管词频有波动,但整体是呈上升趋势的关键词,例如"Absorptive capacity""Citation analysis""Classification""Indicator"等;第二类是研究热度有所下降的词汇"Algorithm""Bibliometrics""Case study"等;第三类是在所统计的十五年里年度波动较大,整体有升有降,例如"Cluster analysis""Data mining""Consumer trust"等;第四类是在"十三五"期间所涌现出来的热门词,是近 5 年的核心热门内容,例如"Altmetrics""Big data""Social media""Information technology"等,结果如表 3 所示:

表 3 "十三五"期间外文期刊论文高频关键词年均频次(部分)

位次	高频关键词	五年总频次	2016 年	2017 年	2018 年	2019 年	2020 年	年篇均频次
1	Internet	396	19.85	21.73	17.47	15.46	13.19	17.73
2	Model	307	13.65	14.86	10.64	14.06	17.45	13.74
3	Social media	253	8.93	10.86	10.24	13.91	11.49	11.32
4	System	152	5.46	6.21	6.43	8.04	7.66	6.80
5	Information technology	124	5.71	5.10	6.43	4.64	6.81	5.55
6	Knowledge	124	3.72	6.65	5.82	4.79	8.09	5.55
7	Satisfaction	86	3.72	3.10	3.41	4.95	3.40	3.85
8	Citation	61	1.24	3.55	4.02	2.32	2.13	2.73
9	Community	59	3.47	3.55	2.01	1.85	2.98	2.64
10	Ecommerce	53	3.47	1.55	2.21	2.16	2.98	2.37
11	Bibliometrics	52	2.98	2.66	2.41	2.01	1.28	2.33
12	Citation analysis	51	3.23	2.22	2.41	1.85	1.70	2.28
13	Big data	50	1.99	1.55	2.21	2.47	3.40	2.24
14	Library	45	1.74	0.00	3.21	2.47	2.55	2.01
15	User acceptance	44	0.00	3.33	2.41	1.39	3.40	1.97
16	Virtual community	44	2.23	2.44	1.81	1.70	1.70	1.97
17	Index	38	1.24	2.44	1.81	1.39	1.70	1.70
18	Indicator	37	1.74	2.66	2.21	1.08	0.00	1.66
19	Academic library	36	2.98	0.00	2.01	1.70	1.28	1.61
20	Algorithm	33	2.48	1.55	1.61	1.24	0.00	1.48

注:为减少小数点位数、便于观察,我们将关键词的相对频次都扩大了 100 倍,即网格中数值 = 相对词频值 ×100

基于表 3 与图 5、图 6、图 7,我们可以得出以下几点结论:

第一,呈现上升态势的关键词,在 2006 - 2015 年里一直缓慢上升(部分关键词有波动,以整体趋势为准),"十三五"期间稳中有进,可见该关键词是近一段时间的研究热点,也是未来一段时间内图情学科的核心热门内容,例如"Citation analysis""Classification""Community""Indicator"等,对于引文分析的研究主要包括以下三种:引文数量分析、引文间网状关系研究、引文反映的主题相关性研究,如测定学科的影响和重要性、探讨研究文献的老化规律等,涵盖年代、引文量、国别、集中或离散规律,技术日趋完善,应用不断扩大。由于引文分析不要求其它先决条件和辅助条件,未来一段时间也将成为普遍关注的方向。此外,我国学者还热衷于传统评价指标之外的新期刊评价指标的研究,如 g-index、h-index、e-index 等,这也是这一时期的研究热点。

第二,呈现下降态势的关键词例如"Accessibility""Algorithm""Case study""Customer satisfaction"等,这代表研究热点在萎缩,但部分关键词出现频次减少并

不代表热度降低，以“Algorithm”为例，由于大数据时代的到来，信息爆发增长，基于算法的推荐则可以解决这一问题，目前学者对于算法的研究不仅仅是以算法为对象，更多是基于算法进行模型构建分析、推荐系统改进、人工智能深度学习等；另一类热度下降的关键词以“Bibliometrics”为例，集数学、统计学、文献学为一体，注重量化综合性知识，对文献定量化的研究可以回溯到 20 世纪初[16]，并且在情报学内部的逻辑结构中，文献计量学渐居核心地位[17]，但是国内外学者对于文献计量学的研究已经较为全面，文献计量学也形成较为系统的体系，加上替代计量学的出现，学者们的关注热点有所转变。

第三，呈现起伏动态的关键词如“Acceptance”“Consumer trust”“Data mining”“Digital library”，在所统计的年度区间里没有规律性，研究热度受政策、时代背景等多方面因素影响。分析其原因，如“Digital library”，数字图书馆起源于二十世纪九十年代[18]，是一种借用互联网的数字技术，对图书馆书籍进行管理，为用户提供信息的服务系统[19]，在 2016 年时对数字图书馆的评价体系、软件系统研究进入一个理论性争议较大的阶段，因此在这一时期发文量较多，而后，由于数字图书馆的资源建设复杂性与技术性要求高，知识产权侵权案增加，都或多或少制约了对这一主题的研究，以致在所调查的年份里呈现上下起伏的波动。

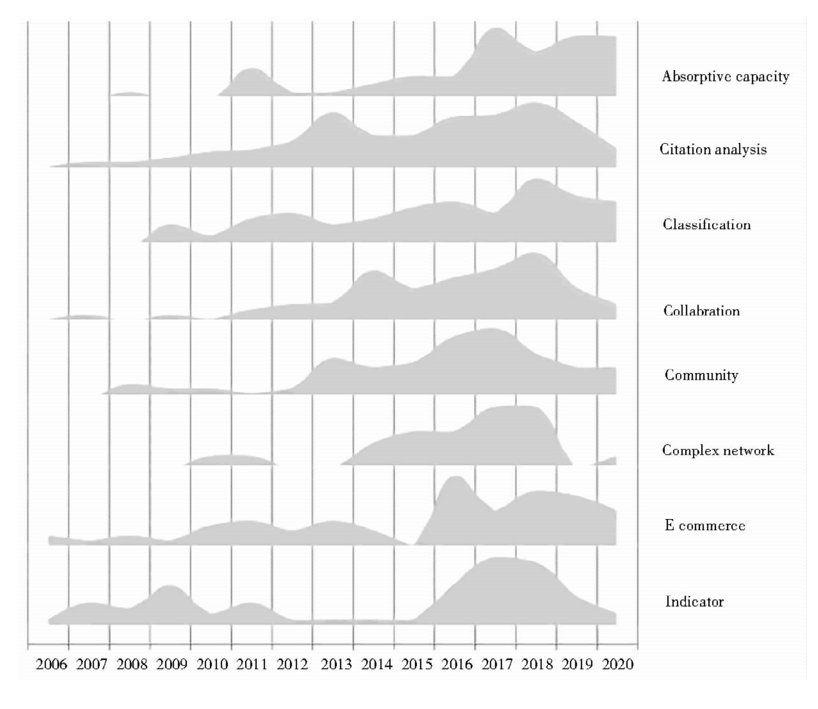


图 5 热度上升词汇

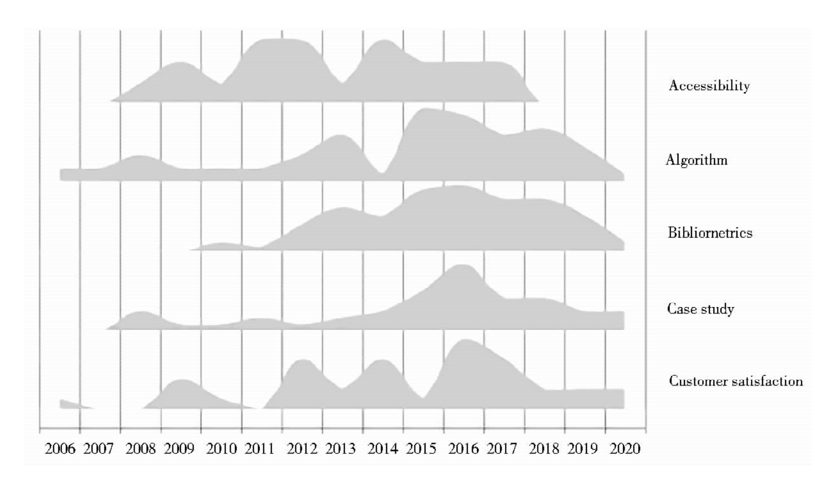


图 6 热度下降词汇

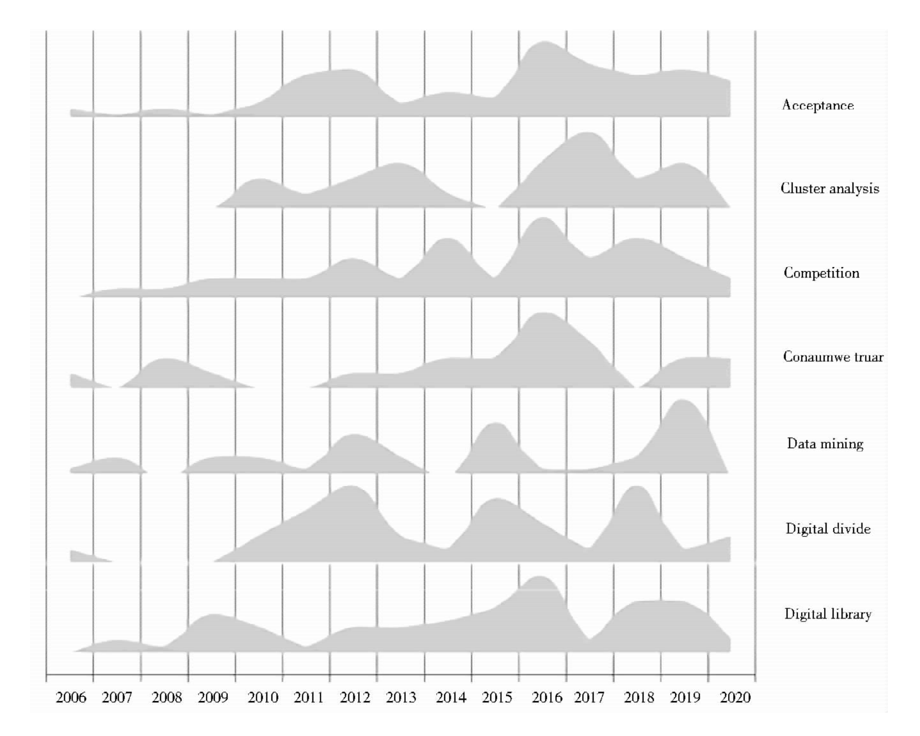


图 7　呈波动态势的词汇

为了更直观表明高频关键词之间的关系，本文还根据相关性对高频关键词进行聚类，对"十三五"期间的高频关键词进行共词检索，得到 54 × 54 的原始共词矩阵，如表 4 所示：

表 4　高频词共词矩阵(部分)

	Social Media	Bibliometrics	Citation Analysis	Case Study	Trust	Knowledge Sharing	Academic Libraries	Machine Learning	Altmetrics	Big Data
Social Media	0	0	1	0	3	1	2	2	2	1
Bibliometrics	0	0	8	0	0	0	0	1	2	0
Citation Analysis	1	8	0	0	0	0	0	0	5	0
Case Study	0	0	0	0	0	0	0	0	0	0
Trust	3	0	0	0	0	3	0	0	0	0
Knowledge Sharing	1	0	0	0	3	0	0	0	0	0
Academic Libraries	2	0	0	0	0	0	0	0	0	2
Machine Learning	2	1	0	0	0	0	0	0	0	2
Altmetrics	2	2	5	0	0	0	0	0	0	0
Big Data	1	0	0	0	0	0	2	2	0	0

在计量分析的过程中，所统计的高频关键词共现频次的多少会受到其他因素的影响，如各自词频的大小，因此，为准确揭示关键词间的共现关系，减少统计时的误差，消除共词差异在数据分析中的影响，本文将高频词共词矩阵转化为相异矩阵并利用 SPSS 软件中的系统聚类法，对共词相异矩阵进行聚类分析，聚类方法上选择组间距离法，点间距离的测度方法选择欧氏距离，见表 5。

在 SPSS 的聚类分析中，以高频关键词的相似性对其进行分类，相似性的度量一般通过变量间的距离来测算，聚类谱系树状图见图 8，从聚类结果看：

第一，关于学科理论与事业的研究。对于科学基础理论的研究不断深入，从图情学科理论基础、研究对象、学科结构与体系等内容多方面进行探讨。我国图情学科的基本研究领域、重要研究热点依然稳定，如"高校图书馆""聚类分析"等相关研究依然是学者的关注重点。但是，随着信息技术的不断发展，图情研究领域也呈现一些新的特征，如情报工作信息化、信息技术创新化、信息检索数字化、用户服务个性化、基础理论的人文化等。由此可见，我国图情学科的研究内容与时俱进，辐射范围较广，发展前景广阔。

第二，关于数字资源的研究。数字资源建设的研

表5 高频词相异矩阵(部分)

	Social Media	Bibliometrics	Citation Analysis	Case Study	Trust	Knowledge Sharing	Academic Libraries	Machine Learning	Altmetrics	Big Data
Social Media	0	1	0.984	1	0.947	0.980	0.959	0.956	0.955	0.977
Bibliometrics	1	0	0.814	1	1	1	1	0.968	0.935	1
Citation Analysis	0.984	0.814	0	1	1	1	1	1	0.811	1
Case Study	1	1	1	0	1	1	1	1	1	1
Trust	0.947	1	1	1	0	0.888	1	1	1	1
Knowledge Sharing	0.980	1	1	1	0.888	0	1	1	1	1
Academic Libraries	0.959	1	1	1	1	1	0	1	1	0.904
Machine Learning	0.956	0.968	1	1	1	1	1	0	1	0.897
Altmetrics	0.955	0.935	0.811	1	1	1	1	1	0	1
Big Data	0.977	1	1	1	1	1	0.904	0.897	1	0

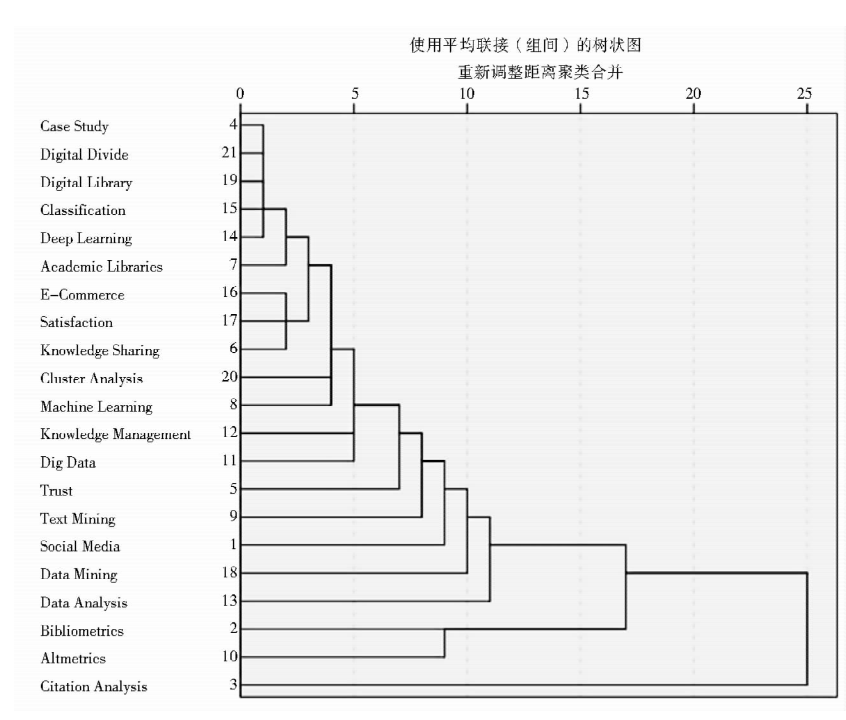


图8 “十三五”期间高频关键词SPSS聚类树状

究从资源的数字化，到资源整合使用方式，再到数据处理的新方式，数字时代的用户需求与服务方式由简单到复杂、由平面到立体、由独占到共享的方向发展，如“Digital library”“Digital divide”“Data mining”。目前学者对该主题的关注点主要包括：数字图书馆的用户隐私保护、与社交媒体的融合、使用者与数字资源间的交互评估、利用现有的大数据挖掘数字资源在数据、技术、服务和用户方面的问题和潜力等，这也是近年来数字资源建设研究的亮点和热点。

第三，关于计量学与期刊评价的研究。计量学和期刊评价作为图情档学科的增长点，以学术期刊为研究对象，使学科知识体系大大延伸。虽然量化学术评价体制引起众多批判，但究其原因在于科研主体机构对评价工具的不当使用。对于评价指标的选取研究，目前学界从单指标评价向多指标评价过渡，单个文献计量指标尽管丰富了评价的选择，但在指标的适用性上却没有一种普遍的检验方式，而多指标的综合评价则主要衡量指标之间的关系和各种计量指标的拟合程度。在评价方法的选择上，基于定量评价的客观性与可操作性，国内学者主要采用这一方法，无论是标准TOPSIS、因子理想解或是效用函数合成，对于综合评价方法的选取一直的学者的研究方向。

第四，关于信息组织与检索的研究。信息技术更加多元，个性化服务引起更多关注。P. Yue[20]提出了预测计算强度的机器学习方法，基于人工智能机器学习改进传统计算；W. Zhang[21]通过利用Word2Vec、余弦词向量构建情感二叉树和基于依赖的情感计算规则，通过大数据分析研究适用于深度学习方法的框架，

探索信息发布策略的影响机制；在个性化服务上，主要包括用户制定的个性化需求和个性化信息检索，涉及云计算、数据挖掘等技术维度关键词和满意度、行为意向等人文维度关键词，用户分类研究是个性化服务研究的重点，这也关系到个性化服务所呈现的效果。

3.5　研究趋势

统计近5年的高频关键词的，有部分关键词在近5年突然涌现成为热门关键词，而此前十年里相关研究涉猎较少，见图9，以下对这些关键词展开分析。

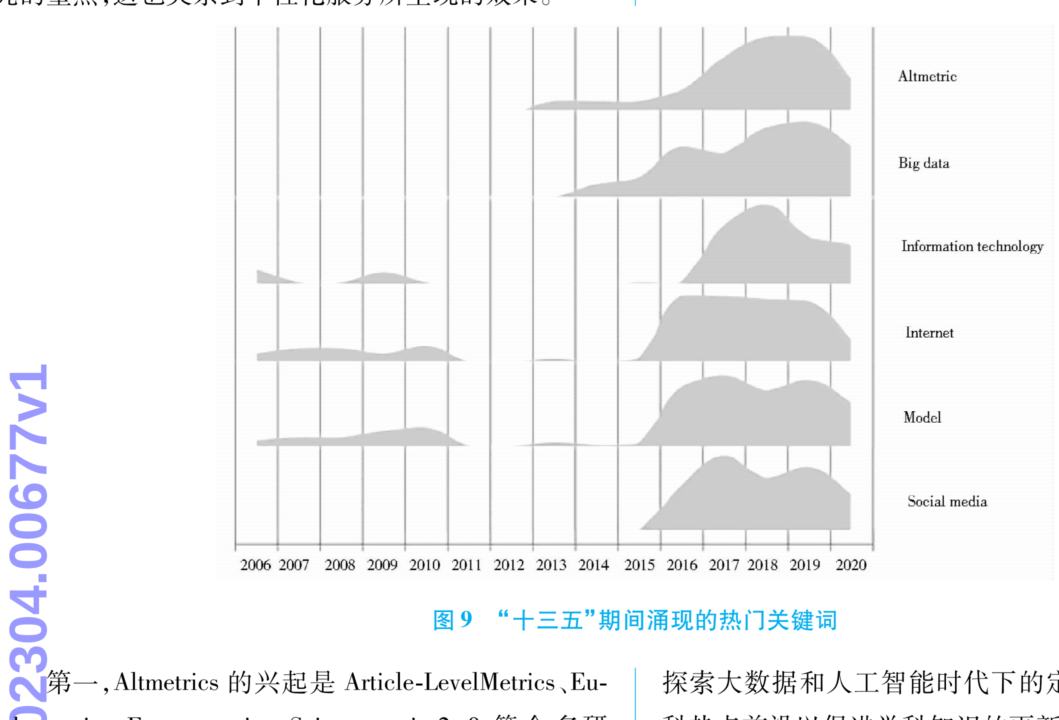


图9　"十三五"期间涌现的热门关键词

第一，Altmetrics 的兴起是 Article-LevelMetrics、Eurekometrics、Erevnametrics、Scientometrics2.0 等众多研究的合流[22]，与科学交流的网络化密切相关，既是提高科学交流效率的需要，也是网络时代学术交流偏好变化的产物。2010 年，Priem 正式提出"Altmetrics"一词，国内学者多将其翻译为"替代计量学"。替代计量学作为新兴的计量研究，重视通过基于社会网络、社交媒体的计量指标来全面评价学术成果的影响力。目前关于替代计量学的研究还处于起步阶段，关注的焦点主要是 Almetrics 指标的内涵、数据来源等，Almetrics 要形成完整且系统的方法体系还要很长一段路，在未来一段时间内这也将是研究的重点方向。

第二，大数据主要兴起于 2012 年[23]，我国学者对于大数据的研究涉及知识管理、数据挖掘、智能决策、移动终端、风险识别、云计算、云服务等。如利用大数据从基于区块链的模型开发信息锚定法、密钥分配法、信息加密算法和防伪算法等核心方法和算法，通过大数据建设集成信息架构与业务模型的智能社区等。大数据是信息社会纵深发展的产物，也是信息社会走向智慧社会的重要标志，其引起的技术剧变所带来的社会变革对图情学科研究的导向作用愈发明显，图情学科作为一门与信息技术息息相关的学科，当前正不断探索大数据和人工智能时代下的定位与发展，追逐学科热点前沿以促进学科知识的更新。

第三，模型构建上主要包括原有理论模型如"信息检索模型"、"知识科学理论模型"和基于研究建立的模型如"模型仿真"、"语义模型"等，利用智能化的技术和理论化的工具建立虚拟学术社区，通过计算、模拟提供新型数字化服务，就目前的研究现状进行预测，模型构建未来仍将是研究热点。

第四，图情领域信息技术的研究主要体现在对信息的处理，利用信息技术支持行业发展、探究信息需求与获取偏好之间的关系等。随着社会的发展，人们获取信息的方式逐渐多样化，更偏爱于信息技术的使用，互联网技术已经在社会各领域得到广泛应用。互联网背景下图情学科的研究包括互联网思维下图书馆服务的改进、"信息疫情"下公众信息素养的培养、互联网所引领的数字图书馆等。2016 年以来国内基于互联网条件下的相关研究发文量呈逐年递增趋势，并且在当前"互联网+"蓬勃发展的背景下，Web2.0 的发展以及移动互联网使用范围迅速膨胀，Facebook、Twitter 以及 Micro blog 等社交软件的类型逐渐增多，各种类型的社交媒体以完全不同于大众传播的人际传播方式迅速成为信息传播与交流沟通的主要渠道，学术交流也

逐渐应用社交媒体而展开,在社交媒体的视角下,丰富的信息知识扩散和交互手段,促使图书情报领域的信息交流渠道以及发展方向涉及到各个新型的领域,可以预测在未来几年中关于互联网的相关研究还会持续升温。

低频词一般不代表学科的研究热点,但一些低频词却能体现局部研究热点的变化,尤其是一些最新出现的低频词,可能就是学科研究前沿的闪现。为进一步探测学科研究前沿以及学科新的生长点,发掘并预测未来重要的研究发展动向,本文对关键词统计中的低频词进行分析,选取五年总频次出现次数为 5 – 20 的关键词进行分析,筛选出属于人文维度、技术角度和管理角度的相关关键词。

本文通过对近 5 年图书情报学领域研究的热点和前沿进行剖析,发现前 100 个关键词中关于技术的关键词与前人所作热点研究所列关键词较为类似,这也说明对于技术的研究在我国图情领域占据重要地位。国内不少学者认为,技术是发展的动力,图情领域呈现出的信息检索和信息资源的研究、知识和数据挖掘的研究均要通过技术实现,在我国的图情领域研究中,技术化现象比比皆是。但早在 1993 年,蒋永福先生指出在图书馆学这门学科研究中不能忘记人的因素,主张在理论的具体实践领域重视人的主体地位并基于这一主体构建相应理论研究体系。他指出,人这一要素处于核心地位,要立足于从用户的维度开展理论研究,主要目的是弘扬学科人文精神。此外,进入 21 世纪以来,图情档学科主体机构管理面临着众多挑战,包括技术变化的挑战、行为与质量管理的挑战、合作与全球化的挑战,在管理大趋势下,主体机构从资源到服务、从硬件到软件、从流程到应用新技术都提出新要求,因此,本文将低频关键词从人文、技术、管理三个维度进行分析,以求发掘未来重要的研究发展动向。

由于在热点分析上已统计频次大于 20 的关键词,对五年间只出现一次或者两次的关键词进行分析无意义,在此我们对词频在 5 – 20 间的低频词进行分析。选择低频词分析对象的原则有以下 3 点:①在后几年新出现的;②相比词频增加较大的;③相比词频增加虽然不大但却有可能代表潜在研究热点的。据此,我们挑选出 64 个低频词作为分析对象,并根据上文的研究,从人文、技术、管理 3 个维度对其进行分析[24]。人文维度的关键词主要是从人文角度和社会科学角度对图情档学科的理论形成、演变、特征、影响、学者感知意向、个体行为和机制发展进行探讨;技术维度的关键词主要是从技术层面对图情档学科数据的挖掘、演化、传播方面进行数据和模型的处理和分析;管理维度的关键词主要包括图情档学科主体机构各种计划、组织、协调和控制的活动、组织管理体系、管理机构、业务管理、发展规划、工作制度及馆际交流等。

(1)代表人文维度的关键词有 26 个,如表 6 所示:

表 6 属于人文维度的关键词

编号	关键词	频次	编号	关键词	频次
1	User satisfaction	18	14	Information literacy	8
2	Purchase intention	17	15	Happiness index	8
3	Behavioral intention	16	16	Perceived usefulness	8
4	Intrinsic motivation	14	17	Information privacy	8
5	Library spirit	13	18	Public lending right	8
6	Privacy concern	13	19	Core value	7
7	Consumer trust	13	20	Freedom of information	7
8	Creativity	12	21	Interpersonal trust	6
9	Individual difference	10	22	User experience	6
10	Library service	10	23	User needs	6
11	Information fairness	9	24	Psychological ownership	6
12	Humanistic care	9	25	Initial trust	5
13	Dynamic capability	9	26	Library Law	5

在对关键词进行筛选后,统计得出属于人文维度的关键词 26 个,主要围绕用户行为、制度机制、感知有用性几个方面展开:

- 用户行为研究。在所统计的关键词中,用户行为包括“Purchase intention”“Behavioral intention”“Consumer trust”“Individual difference”“User experience”,基于意愿、差异、体验等指标进行分析,从中发现用户行为的规律,并将这些规律与相关理论等相结合,从而发现可能存在的问题,这一种模式也将在未来继续深入发展。

- 制度机制研究。在分析近 5 年低频关键词中发现,对于机制的研究主要体现在两方面,一方面是对传统的图书馆服务、制度、隐私问题、深度学习等方面进行研究,以更好地提升和完善图书馆的社会服务功能;另一方面对用户权利保障的相关研究,如关键词“Privacy concern”“Privacy concern”“Freedom of information”等。结合前文高频关键词如“Information resource”,笔者认为“人文视角下用户权利保障机制”将是图书馆学制度研究的趋势。

- 感知有用性研究。如“User satisfaction”“Con-

sumer trust”“Initial trust”，感知有用性用来探讨外部因素对用户的信念、态度及意向的影响，与态度倾向成正比，这也是影响用户个人信息披露意愿的重要因素，在相关论文中主要体现在对信息系统的价值认可，信息输出通透、系统功能灵活与否上，对这一关键词的研究有助于更好地理解个体信息行为，这也有可能成为未来一段时间研究热点。

（2）代表技术维度的关键词有 25 个，如表 7 所示：

表 7　与技术相关的关键词

编号	关键词	频次	编号	关键词	频次
1	Ontology	19	14	RSS	9
2	Structural equation model	19	15	Artificial intelligence	9
3	Data analysis	18	16	Genetic algorithm	9
4	Network analysis	17	17	Information sharing	8
5	Collaboration network	14	18	Co-citation analysis	8
6	Database	13	19	Spatial data mining	8
7	Open source software	12	20	Web2.0	7
8	Citation network	12	21	Open street map	5
9	Visualization	11	22	Real-time consultation	5
10	IOT	11	23	Knowledge network	5
11	Information retrieval	11	24	Semantic web	5
12	Cloud computing	10	25	Library 2.0	5
13	Recommended system	9			

近 5 年图情学科外文期刊中在技术方面的关键词可以总结为两大类：数据挖掘与基于 web2.0 的信息服务，数据挖掘主要是从大量的数据中搜索隐藏于其中的特殊关系性，通过统计、情报检索、在线分析处理、机器学习等方法来实现；而基于 web2.0 的信息服务如“RSS”“Visualization”“Open source software”“Real-time consultation”等，进入 web2.0 时代后，语义 web、可视化分析、虚拟社区等引起了学界关注，学者们纷纷从社会学角度、商业角度、组织与信息系统角度等多个视角研究如何提供新型的数字化服务，利用其核心技术如 RSS、Tag 拓展网络服务下的信息范围，延伸信息服务，发掘隐性知识。就目前的研究现状进行预测，未来这两方面将成为重要研究主题。

（3）代表管理维度的关键词有 13 个，见表 8。

根据统计，图情档外文期刊近 5 年在管理维度上的研究偏向于图书馆的交互、定位、技术创新等，图书馆学研究也引入了一些管理学的理念，如“Business outsourcing”“Crisis management”“Library management”“Human resource Management”等，且更为注重绩效管理，如“Library evaluation”“Quality Evaluation”等。分析

表 8　与管理相关的关键词

编号	关键词	频次	编号	关键词	频次
1	Library evaluation	16	8	Supply chain management	7
2	Business outsourcing	15	9	Market orientation	6
3	Innovation performance	11	10	Quality evaluation	6
4	Job performance	10	11	Organizational change	5
5	Service innovation	9	12	Technology adoption	5
6	Library management	7	13	Human resource Management	5
7	Crisis management	7			

其原因，一方面随着信息化时代的不断发展，高校图书馆利用信息资源、尖端技术和科学方法为为企业开展特色服务，同时涵盖对企业图书馆的信息汇集、加工处理与生产信息并存的研究；另一方面当前信息社会的挑战、市场经济的驱动、网络化、数字化的冲击以及环境的变化和社会的加速进步，都促使图书馆在管理上不断变革，相关管理一直是学科的一个重要关注点，在未来一段时间关注度也将呈现上升的趋势。

4　结语

本文以高影响力外文期刊所刊载的 2 270 篇研究论文及其题录数据作为研究样本，利用信息计量研究方法，从多方面对近 5 年 LIS 研究领域外文期刊的研究热点和研究前沿领域进行分析与展示，借助相关文献分析软件对近 5 年来图书情报学研究现状、热点和前沿进行了可视化分析。通过相关的统计数据，分别从研究现状、高影响力作者、研究热点和研究前沿四个方面进行了分析，并得出以下研究结论：

首先，从发文量和被引频次两个角度分析了近年来图书情报学研究领域的高影响力作者，这些研究者在图书情报学研究领域表现得极为活跃，不仅推动了图书情报学学科的发展和进步，而且在一定程度上引领着未来图书情报学的研究方向和趋势。

其次，通过分析高被引文章，揭示“十三五”时期的关注热点主要集中在基础理论、前沿技术和现实社会关系三方面。

再次，通过高频关键词的年度变化分析，总结三类呈现不同态势的关键词，分别为图情领域的核心热门词、持续热点词以及研究热度呈起伏状态的关键词。

最后通过研究前沿的分析，我们发现随着信息技术的不断发展和信息环境的不断变化，图情学科研究主题也会随之发生转变，由传统信息服务研究向知识服务研究转变，由知识组织与管理相关研究向知识组

织体系构建等研究主题进一步延伸和细化,而注重理论和实践相结合也将日渐成为图书情报与档案管理学科研究的重要趋势。

参考文献:

[1] WHITE H D, MCCAIN K W. Visualizing a discipline: An author co-citation analysis of information science[J]. Journal of the Association for Information Ence & Technology, 2010, 49(4):327 - 355.

[2] 邱均平, 温芳芳. 近五年来图书情报学研究热点与前沿的可视化分析——基于 13 种高影响力外文源刊的计量研究[J]. 中国图书馆学报, 2011, 37(2):51 - 60.

[3] 邱均平, 马瑞敏. 基于 CSSCI 的图书馆、情报与档案管理一级学科文献计量评价研究[J]. 中国图书馆学报, 2006, 32(1): 24 - 29.

[4] 胡浪. 2012 - 2016 年国内图书情报学研究现状及热点探析——基于 13 种中文图情核心期刊的计量分析[J]. 图书馆工作与研究, 2018(2):67 - 76.

[5] D·普赖斯, 张季娅. 洛特卡定律与普赖斯定律[J]. 科学学与科学技术管理, 1984(9):17 - 22.

[6] 百度百科. 普赖斯定律[EB/OL]. [2020 - 07 - 31]. http://baike.baidu.com/view/2288612.htm.

[7] PAO M L. Lotka's Law: a testing procedure[J]. Information processing & management, 1985, 21(4):305 - 320.

[8] NAMBISAN S, LYYTINEN K, MAJCHRZAK A, et al. Digital innovation management: reinventing innovation management research in a digital world[J]. Mis quarterly, 2017, 41(1):223 - 238.

[9] HOU J, YANG Y, CHEN C. Emerging trends and new developments in information science: a document co-citation analysis (2009 - 2016)[J]. Scientometrics, 2018,115(2):869 - 892.

[10] HASHEM I A T, CHANG V, ANUAR N B, et al. The role of big data in smart city[J]. International journal of information management, 2016, 36(5):748 - 758.

[11] SHIAU W L, CHAU P Y K. Understanding behavioral intention to use a cloud computing classroom: a multiple model comparison approach[J]. Information & management, 2016, 53(3):355 - 365.

[12] YAO Y, LI X, LIU X, et al. Sensing spatial distribution of urban land use by integrating points-of-interest and Google Word2Vec model[J]. International journal of geographical information science, 2016, 31(4):825 - 848.

[13] REHMAN M H U, CHANG V. Big data reduction framework for value creation in sustainable enterprises [J]. International journal of information management. 2016, 36 (6): 917 - 928.

[14] LI L, SU F, ZHANG W, et al. Digital transformation by SME entrepreneurs: a capability perspective[J]. Information systems journal, 2018, 28(6):1129 - 1157.

[15] 孙清兰. 高频词与低频词的界分及词频估算法[J]. 中国图书馆学报, 1992(2):78 - 81.

[16] 赵蓉英, 许丽敏. 文献计量学发展演进与研究前沿的知识图谱探析[J]. 中国图书馆学报, 2010(5):60 - 68.

[17] 屈宏明. 用文献计量学方法对网络环境下馆藏文献的评价[J]. 现代情报, 2008, 28(6):27 - 29.

[18] 叶晴. 数字时代研究型高校图书馆空间变革研究[D]. 杭州:浙江大学,2019.

[19] 李洁. 数据驱动下数字图书馆知识发现服务创新模式与策略研究[D]. 长春:吉林大学,2019.

[20] YUE P, GAO F, SHANGGUAN B, et al. A machine learning approach for predicting computational intensity and domain decomposition in parallel geoprocessing [J]. International journal of geographical information science, 2020, 34(11): 2243 - 2274.

[21] ZHANG W, WANG M, ZHU Y C. Does government information release really matter in regulating contagion-evolution of negative emotion during public emergencies? from the perspective of cognitive big data analytics[J]. International journal of information management, 2020, 50(2):498 - 514.

[22] 邱均平, 余厚强. 替代计量学的提出过程与研究进展[J]. 图书情报工作, 2013, 57(19):5 - 12.

[23] 孟小峰, 慈祥. 大数据管理:概念、技术与挑战[J]. 计算机研究与发展, 2013, 50(1):146 - 169.

[24] 邱均平,丁敬达. 1999 - 2008 年我国图书馆学研究的实证分析(下)[J]. 中国图书馆学报,2009,35(6):79 - 87,118.

[25] 赵蓉英, 赵洝吟, 陈必坤. 我国"图书馆,情报与档案管理"学科的研究现状分析——以 2001 - 2012 年国家基金项目为视角[J]. 情报杂志, 2013, 32(7):106 - 112.

[26] 赵蓉英, 王旭. 近 5 年国内外情报学研究态势比较分析[J]. 情报科学, 2018, 36(7):3 - 9.

[27] 刘奕杉, 王玉琳, 李明鑫. 词频分析法中高频词阈值界定方法适用性的实证分析[J]. 数字图书馆论坛, 2017(9):42 - 49.

[28] 张立, 尹萍. 我国图情领域近十年国际发文研究热点和前沿剖析[J]. 情报科学, 2020, 38(5):82 - 89.

作者贡献说明:

赵蓉英:论文选题、框架设计、论文指导;
朱伟杰:数据分析、论文撰写、论文修改;
常茹茹:数据预处理;
段香廷:数据下载。

赵蓉英，朱伟杰，常茹茹，等."十三五"图情档学科进展:外文期刊论文解析[J]. 图书情报工作,2021,65(5):26－39.

Progress of Library and Information Science and Archival Science During the 13th Five-year Plan: Analysis of the Articles in Foreign Journals

Zhao Rongying[1,2,3] Zhu Weijie[1,3] Chang Ruru[1,3] Duan Xiangting[1,3]

[1] Research Center for Chinese Science Evaluation, Wuhan University, Wuhan 430072

[2] Center for Studies of Information Resources, Wuhan University, Wuhan 430072

[3] School of Information Management, Wuhan University, Wuhan 430072

Abstract: [**Purpose/significance**] The research on the international publication, research hotspots and periodical characteristics of the Library and Information Science is helpful to promote the development of the subject and enhance its international discourse power. [**Method/process**] In this paper, we analyzed the key words frequency of papers published in high-impact foreign journals included by SCI and SSCI, and used HistCite, CiteSpace, SPSS and other tools to analyze the development of the chart in the past five years (2016－2020) from five aspects including current research status, high-impact scholars, highly cited papers, research hotspots, and cutting-edge trends. Through the annual change analysis of high-frequency keywords, four keywords with different trends were summarized, and the low-frequency keywords are perspectived. The development themes of humanity, technology and management were obtained. [**Result/conclusion**] With the continuous development of information technology and changes in the information environment, the themes of the Library and Information Science have been further extended and refined from traditional information services to knowledge services, and from research related to knowledge organization and management to the construction of knowledge organization systems, and the combination of theory and practice has become an important trend in the research of Library and Information Science.

Keywords: 13th Five-Year Plan library and information science discipline progress journal articles in foreign languages word frequency analysis